Infant Behavior and Development 57 (2019) 101346

Contents lists available at ScienceDirect

Infant Behavior and Development

journal homepage: www.elsevier.com/locate/inbede

Full length article

Rhythmic structure facilitates learning from auditory input in newborn infants

Emma Suppanen[a,*], Minna Huotilainen[a,b], Sari Ylinen[a,b]

[a] Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland
[b] Cicero Learning, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

ARTICLE INFO

Keywords:
Auditory processing
Electroencephalography (EEG)
Event-related potential (ERP)
Predictive coding
Music
Nursery rhyme
Language development

ABSTRACT

Rhythm and metrical regularities are fundamental properties of music and poetry - and all of those are used in the interaction between infants and their parents. Music and rhythm perception have been shown to support auditory and language skills. Here we compare newborn infants' learning from a song, a nursery rhyme, and normal speech for the first time in the same study. Infants' electrophysiological brain responses revealed that the nursery rhyme condition facilitated learning from auditory input, and thus led to successful detection of deviations. These findings suggest that coincidence of prosodic cue patterns and to-be-learned items is more important than the format of the input. Overall, the present results support the view that rhythm is likely to create a template for future events, which allows auditory system to predict prospective input and thus facilitates language development.

1. Introduction

Across cultures, caregivers and their infants use music to interact: for example, singing play songs and lullabies, reciting nursery rhymes, and rocking infants in time with music (Fernald & Kuhl, 1987; Obermeier et al., 2013; Papousek, 1996; Shannon, 2006; Trehub & Schellenberg, 1995; Wallin et al., 2001). Previous research has shown that this behavior results in beneficial socioemotional outcomes (Cirelli, Einarson, & Trainor, 2014; Kivijärvi et al., 2001), and supports early language development.

Many studies have shown the benefits of both formal and informal musical activities to auditory, linguistic and literacy skills in children (e.g. Bergman Nutley, Darki, & Klingberg, 2014; Chobert, François, Velay, & Besson, 2012; François & Schön, 2011; François, Chobert, Besson, & Schön, 2012; Kraus et al., 2014; Linnavalli, Putkinen, Lipsanen, Huotilainen, & Tervaniemi, 2018; Moritz, Yampolsky, Papadelis, Thomson, & Wolf, 2013; Overy, 2003; Tallal & Gaab, 2006; Torppa et al., 2014; for a review, see Virtala & Partanen, 2018). For example, music training allows children to detect small pitch changes (Besson, Schön, Moreno, Santos, & Magne, 2007), and children's music experience has been found to promote verbal memory (Ho, Cheung, & Chan, 2003), reading skills (Anvari, Trainor, Woodside, & Levy, 2002; Corrigall & Trainor, 2011), vocabulary (Forgeard, Winner, Norton, & Schlaug, 2008; Linnavalli et al., 2018), phoneme awareness (Anvari et al., 2002), and detection of prosody (Magne, Schön, & Besson, 2006). Music exposure changes the auditory cortical processing (Partanen, Kujala, Tervaniemi, & Huotilainen, 2013; Trainor, Lee, & Bosnyak, 2011) and increases the number of preverbal communicative gestures in 12-month-old children (Gerry, Unrau, & Trainor, 2012). Music captures and sustains infant attention, and infants are more engaged when listening to maternal singing than to maternal speech (Nakata & Trehub, 2004). In addition, musical activities seem to modify the phase of attention fluctuations (Jones, 1976) to

* Corresponding author at: Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, P.O. Box 9, FIN-00014, Helsinki, Finland.

E-mail address: emma.suppanen@helsinki.fi (E. Suppanen).

https://doi.org/10.1016/j.infbeh.2019.101346
Received 22 May 2019; Received in revised form 31 July 2019; Accepted 1 August 2019

the periodicities in the music, thus improving rapid temporal auditory processing and enhancing neural encoding of speech (Kraus & Chandrasekaran, 2010; Schön & Tillmann, 2015; Zhao & Kuhl, 2016).

There have been fewer studies that investigate solely rhythm without the music accompaniment, yet the effects of rhythm have been observed from infancy to adulthood. Metric regularity of rhythmic speech and movement might have beneficial effects on speech processing. It has been found to facilitate word segmentation in 9-month-old infants (Curtin, Mintz, & Christiansen, 2005) and in adults (Leong & Goswami, 2015; Rothermich, Schmidt-Kassow, Schwartze, & Kotz, 2010; Schmidt-Kassow & Kotz, 2009; Schön et al., 2008). Rhythm also helps lexico-semantic integration (Rothermich, Schmidt-Kassow, & Kotz, 2012) and increases aesthetic appreciation and emotional processing (Obermeier et al., 2013) in adults. Rhythm perception is associated with foreign language learning in adults (Bhatara, Yeung, & Nazzi, 2015), and grammar skills in 5–7 year old children (Gordon et al., 2015). Moreover, children with developmental dyslexia have been found to have poor rhythmic perception (2011b, Flaugnacco et al., 2014; Goswami, 2011a; Goswami et al., 2002; Huss, Verney, Fosker, Mead, & Goswami, 2011). Goswami (2011, 2018) has suggested that rhythmic auditory input causes phase synchronization in oscillations of auditory cortex, and thus results in a more efficient neural sampling of the auditory signal.

Bimodal auditory and visual stimuli in synchrony are highly salient, and deploy infants' attention efficiently benefiting their perceptual processing and memory (Bahrick & Lickliter, 2000; Lewkowicz & Lickliter, 2013; Lewkowicz, 2000; Reynolds, Bahrick, Lickliter, & Guy, 2014). Music (Lebedeva & Kuhl, 2010; Peterson & Thaut, 2007; Schön et al., 2008; Thaut, 2005) and rhythm with or without music (Purnell-Webb & Speelman, 2008) have been found to enhance learning and memory in adults. This is because the structure of music and rhythms are likely to create a template for future events. Such structure helps adaptive auditory system to pick up sound regularities in a predictive manner (Rothermich et al., 2010; Schmuckler & Boltz, 1994; Winkler, Háden, Ladinig, Sziller, & Honing, 2009).

Infants can detect pitch and rhythm changes at birth (Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008; Telkemeyer et al., 2011; Trehub & Hannon, 2006) and even before that as near-term fetuses (Graniere-Deferre, Ribeiro, Jacquet, & Bassereau, 2011; Lecanuet, Graniere-Deferre, Jacquet, & DeCasper, 2000; Partanen, Kujala, Näätänen et al., 2013). Furthermore, infants are sensitive to prosodic rhythm, and can discriminate between languages representing different rhythm classes – even when the phonemic structure of speech is degraded (Nazzi, Bertoncini, & Mehler, 1998). Thus, it seems that infants are already sensitive to prosodic information carried by slow-varying patterns of spectral and amplitude modulation.

Infants have been found to be able to extract and predict auditory regularities from sound frequency, duration, timbre (Háden, Németh, Török, & Winkler, 2015; Ruusuvirta, Huotilainen, Fellman, & Näätänen, 2009; Trainor, 2012) and stream of syllables (Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009). François et al. (2017) found that newborn infants were able to detect statistical violations in pseudo-words in melodically enriched, but not in flat speech streams. They also showed that neonatal brain responses to deviations in melodically enriched speech correlated with expressive vocabulary at 18 months. Furthermore, newborn infants can extract temporal predictions from rhythmical regularities in the auditory input (Winkler et al., 2009). However, little is known about how neonates encode continuous speech, how much they are able to learn from it, and whether music or rhythmic speech helps them to do so.

1.1. Current study

The aim of this study was to explore whether music and rhythm can be observed to facilitate learning from auditory input in newborn infants. To this end, we presented natural continuous stimuli with similar semantic content in the form of a nursery rhyme, a song and prose to infants and recorded their auditory event-related potentials (ERPs). We then probed the learning effects by introducing infrequent changes of vowel, word, pitch, and intensity changes to speech excerpts (similar to the optimal paradigm; Näätänen, Pakarinen, Rinne, & Takegata, 2004) and by recording the neural responses to them. In line with predictive coding theory (Rao & Ballard, 1999, 2010, Emberson, Richards, & Aslin, 2015; Friston, 2005; Friston, Kilner, & Harrison, 2006; Kayhan, Meyer, O'Reilly, Hunnius, & Bekkering, 2019; Kouider et al., 2015; Pickering & Garrod, 2013), we expected that sensory stimulation results in attempts to generate an internal model of its regularities in the infant brain. When an incoming stimulus deviates from the internal model, a quick and robust brain response is elicited. This response, known as prediction error, can be thought as a feedback signal to the higher levels of neural network that adjusts the neural representations of the internal model to minimize future surprises (Friston et al., 2006; Winkler, 2007; see also Näätänen, Paavilainen, Rinne, & Alho, 2007, for the mismatch negativity (MMN) elicited in an oddball paradigm). In the case of infrequent changes in our stimuli, possible prediction error brain responses would reflect learning, because without any learning from the long continuous stimuli, the infants' internal model could not predict incoming stimulation (for predictions in speech processing and recognition, see DeLong, Urbach, & Kutas, 2005; Gagnepain, Henson, & Davis, 2012; Ylinen et al., 2016; Ylinen, Bosseler, Junttila, & Huotilainen, 2017). We hypothesized that the rhythm present in both nursery rhyme and song, and the melody present in song would enhance the extraction of predictions from auditory input and thus facilitate learning. We expected this to result in the strengthening of the prediction error response, which in this case reflected infants' learning as a result of exposure.

2. Methods

2.1. Participants

Participants (N = 21; 8 boys, 13 girls) were healthy newborn infants born into Finnish-speaking families. The mothers or both

Table 1
Participant details (N = 21).

	Age (days)	Gestational age (weeks + days)	Weight (g)	Length (cm)	5 minute Apgar score[a]
Average	2.0 (SD 1.15)	39 + 4	3511	50.5	9.8
Range	0.5–5	37 + 3–41 + 6	2510–4835	47–56	8–10

[a] Apgar (**A**ppearance, **P**ulse, **G**rimace, **A**ctivity, **R**espiration, range 0–10) is used to assess the health of the newborn.

parents of the newborn gave written informed consent to participate in the study. Electroencephalography (EEG) measurements were conducted in Jorvi hospital in Espoo, Finland. The study was approved by the Ethics Committee for Paediatrics, Adolescent Medicine and Psychiatry, Hospital District of Helsinki and Uusimaa. Initially 24 newborns participated in the study, but data of three participants were excluded due to technical problems or not reaching the criterion of 40 epochs. More details about participants can be found in Table 1.

2.2. Study paradigm and stimuli

A Finnish version of a well-known nursery rhyme Simple Simon (Kunnas, 1997) was recorded by a female native speaker of Finnish using three different conditions: she read two verses as metrically-spoken nursery rhyme, sang two verses to a melody (see Fig. 1), and read two verses in ordinary speech. To elicit as naturalistic stimuli as possible, the speaker was not given specific instructions with respect to acoustics, and thus the conditions differ by their duration, intensity, pitch and rhythmic structure (see detail in Table 4). However, she was aware that the stimuli would be directed to infants, and was given written instructions to read the stimuli as if she was reading to a toddler.

The ordinary speech was modified from the original nursery rhyme. The modifications of this prose version included converting the atypical word order of the nursery rhyme into normal one, and adding modifiers (genitive attributes) and clarifying words, the meanings of which were implicit in the original nursery rhyme but which had been omitted from the nursery rhyme because of the meter (see Table 2).

The nursery rhyme used in the experiment had a regular rhythmic structure (trochee) that was characterized by regular alteration of stressed and unstressed syllables that were cued by pitch and intensity. The song had a regular beat, but as compared to the nursery rhyme that used both pitch and intensity to cue stress, there was smaller acoustic difference between stressed and unstressed syllables (i.e., stressed syllables were less prominent). Rather than to cue stress, pitch was used to convey melody. Ordinary speech was, in turn, characterized by less regular pattern of stressed and unstressed syllables, since stress placement was determined by focus in the sentence (rather than meter or beat).

In order to be able to examine the acoustic structures of nursery rhyme, song and speech in more detail, we utilized a number of acoustic measures. We identified and labelled the locations of vowel-consonant and consonant-vowel boundaries to segment the stimuli into vocalic intervals (n = 169), located between the onset and the offset of a vowel, or a vowel cluster, and consonantal intervals (n = 170), located between the onset and the offset of a consonant, or a cluster of consonants (Ramus, Nespor, & Mehler, 1999). This was done in Praat (Boersma & Weenink, 2018). When boundaries could not be identified by visual inspection of speech waveforms, a boundary was placed at a point where the preceding phone could no longer be heard. The measures we used were (i) the standard deviation of vocalic interval duration (ΔV), (ii) the standard deviation of consonantal interval duration (ΔC), (iii) the proportion of vocalic utterance duration out of total duration multiplied by 100 (%V), (iv) the standard deviation of vocalic interval duration normalized for speech rate variability (VarcoV), (v) the standard deviation of consonantal interval duration normalized for speech rate variability (VarcoC), (vi) the mean durational differences between successive vocalic or consonantal intervals divided by the sum of the same intervals, the normalized pairwise variability index (nPVI), and (vii) normalized pairwise variability index for vocalic intervals (nPVI-V) (Dellwo & Wagner, 2003; Grabe & Low, 2002; Lee, Kitamura, Burnham, & Todd, 2014; Ramus & Mehler, 1999; Ramus et al., 1999; White & Mattys, 2007a; 2007b; Wiget et al., 2010). White and Mattys (2007b) found especially VarcoV useful for discriminating different rhythm topologies: higher VarcoV suggests more salient contrasts between stressed and unstressed syllables (see Fig. 2). In line with Lee et al. (2014) we also utilized the pitch extraction algorithm in Praat (with standard settings:

Fig. 1. The melody to which the nursery rhyme was sung.

Table 2
The modification of the original nursery rhyme into ordinary speech (syllables with word stress in bold). Note the changes in word order and added modifiers.

Nursery rhyme	Ordinary speech	Literal translation
Äidin **ve**siämpäristä **Pek**ka **on**ki **ka**laa.	**Pek**ka **on**ki **ka**laa **äi**din **ve**siämpäristä.	Pekka was angling for a fish in mother's water bucket.
Koukussansa **on**kimato **nau**reskeli **sa**laa.	**On**kimato **nau**reskeli **sa**laa **Pe**kan **kou**kussa.	Angleworm was secretly laughing on Pekka's fishhook.
Pekka **et**si **nok**kosista **si**nivatukoita.	**Pekka etsi sinivatukoita nokkosista.**	Pekka looked for dewberries in nettle bush.
Kädet **pa**loi **pu**naisiksi, **täy**teen **rak**kuloita.	**Pe**kan **kä**det **pa**loivat **pu**naisiksi ja **tu**livat **täy**teen **rak**kuloita.	Pekka's hands burned red and filled with blisters.

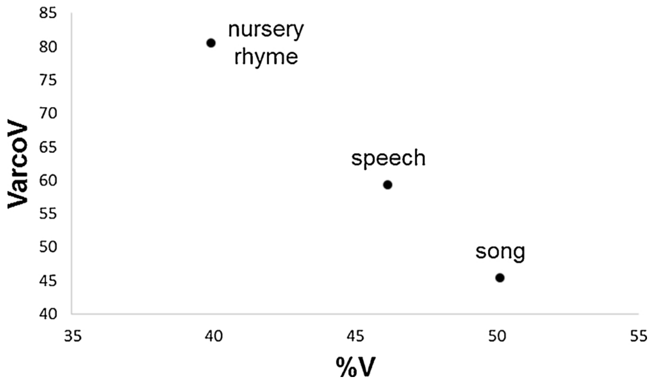

Fig. 2. Distribution of nursery rhyme, song and speech over the plane representing standard deviation of vocalic interval duration normalized for speech rate variability (VarcoV), and the proportion of vocalic utterance duration out of total duration (%V).

75 Hz pitch floor, 500 Hz pitch ceiling, 0.45 voicing threshold) to estimate the median F0 of all vocalic intervals, and then calculated mean pitch in vocalic intervals and the standard deviation of the median pitches of vocalic intervals. These measures and other acoustic details of the stimuli can be found in Table 3.

We were also interested in those prosodic stress patterns in our stimuli which are not revealed by durational measures, such as syllables, words and phrases (for discussion on general problems with durational rhythm metrics, see Arvaniti, 2012). Leong (2012) introduced the idea of using nested amplitude modulations (AM) as a measure of rhythmic information in speech: modulation range 0.9–2.5 Hz represents the stressed syllable rate, and 2.5–12 Hz range the syllable rate. The amplitude envelope (AE) of speech varies relatively slowly in time, and the amplitude rises coincide with stressed syllable onset, which occurs approximately twice per second in infant-directed speech (IDS) and nursery rhymes (Goswami, 2018; Leong & Goswami, 2015; Leong, Kalashnikova, Burnham, &

Table 3
Acoustic details of the stimuli. ΔV and ΔC are the standard deviations of vocalic and consonantal interval durations, respectively, %V is the proportion of vocalic utterance duration out of total duration multiplied by 100, VarcoV and VarcoC are the standard deviations of vocalic and consonantal interval durations normalized for speech rate variability, and nPVI and nPVI-V are the normalized pairwise variability index and normalized pairwise variability index for vocalic intervals.

	Nursery rhyme	Song	Speech
Duration of two verses (s)	18.45	18.95	14.20
Mean Intensity (SD), dB	72.5 (15.1)	74.0 (15.2)	71.3 (14.1)
Median Intensity, dB	67.8	73.0	66.9
Mean Pitch (SD), Hz	185.5 (44.5)	242 (25.1)	170 (30.4)
Median Pitch, Hz	164.5	249.7	157.8
ΔV (ms)	91.7	59.7	53.4
ΔC (ms)	92.5	70.9	57.8
%V	39.9	50.1	46.2
VarcoV (s)	0.81	0.45	0.56
VarcoC (s)	0.55	0.47	0.56
nPVI	16.4	14.3	15.6
nPVI-V	12.7	12.2	18.3
Δpitch (vocalic intervals, Hz)	38.3	27.2	45.0
mean pitch (vocalic intervals, Hz)	161.7	240.1	163.1
n (vocalic intervals)	54	55	60
n (consonantal intervals)	55	56	59

Table 4
Rhythmic structure details. Coefficient estimates describe how much 2 Hz rhythm (Goswami, 2018) each condition has. A higher value means that there is more 2 Hz rhythm in the signal. Average temporal distances were estimated by determining prominent peaks in signal intensity and pitch, and then calculating the temporal distances between those peaks and onsets of sentence stress. The lower the value of average temporal distance, the closer the peaks match the sentence stress.

	Coefficient estimates (envelope and 2 Hz)	Coefficient estimates (pitch and 2 Hz)	Average temporal distance (s)
Nursery rhyme	1.87	0.59	3.55
Song	0.96	0.56	3.95
Speech	0.84	0.33	4.42

Goswami, 2014). For example, Leong et al. (2014) used this method to find that dominant rhythmic patterning of IDS occurs at prosodic stress feet whereas for adult-directed speech (ADS) it happens at syllable rate.

2.3. Descriptive analyses of the stimuli

We extracted the amplitude envelopes of our stimulus signals using Hilbert transform and 40 Hz low-pass filter, and compared the envelopes and pitch contours of the three stimulus types with 2-Hz sine wave by calculating generalized linear regression, where either normalized signal envelope or pitch was the dependent variable and analytical (Hilbert transformed) 2-Hz signal the independent variable (see e.g. Parkkonen, Andersson, Hämäläinen, & Hari, 2008). Analytical signal was used to accommodate the phase of the signal. A higher value means that there is more 2 Hz rhythm in the signal. As indicated by Table 4, generalized linear regression analysis suggested that the highest degree of 2 Hz rhythm was found in the nursery rhyme signal. That is, the regularity of the rhythmic structure as indicated by intensity and pitch was highest in the nursery rhyme and lowest in speech.

In addition, we examined the rhythmic regularity of our stimuli signals by computing the autocorrelation functions (ACF) of the nursery rhyme, song and speech signal envelopes that had been filtered into stressed syllable (0.9–2.5 Hz) and syllable (2.5–14 Hz) modulation rates. A Fourier transform was then applied to the ACFs in order to get a periodic power spectrum for each of the nursery rhyme, song and speech ACFs at both modulation rates (Leong, 2012). Fig. 3 shows these spectra. Based on this analysis nursery rhyme condition contained higher periodic power than the other two conditions at stressed syllable modulation rate, but song condition had slightly higher periodicity at the syllable modulation tier.

These six verses were presented as repetitive standard stimuli to the infants for 10.4 min (12 repetitions/verse) during a learning phase. A testing phase (10 repetitions/verse, 8.6 min), which followed immediately after the learning phase, introduced multiple deviations to the syllables with word stress. The possible deviations were the change of the whole word, vowel, sound intensity, or pitch. The deviations were inserted to all types of verses avoiding two simultaneous ones. The order of the verses was counterbalanced across participant groups (4 participants/group). There were 10 differing versions of each verse, and each altered verse included each deviation type twice (see Table 5). There was only one deviation type per altered syllable, however, and the same deviation did not repeat on each syllable more than once. Intensity change was implemented as six dB increase or decrease of sound pressure level in all conditions. Sound frequency deviations were carried out by increasing or decreasing base frequency by 20% in poetry and prose verses, whereas in song verses melody was shifted by two to five whole tones using notes that did not break the original chord progression (see Appendix A).

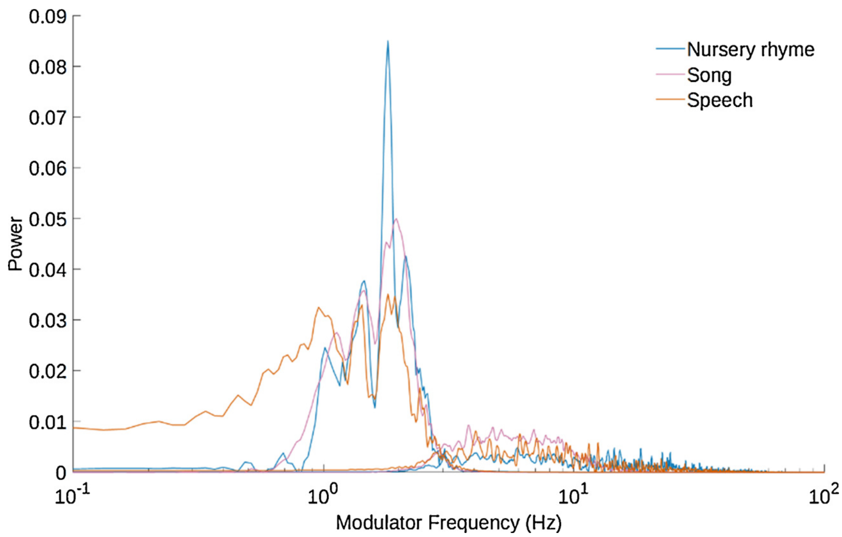

Fig. 3. Periodic power for the autocorrelation functions of both modulation tiers.

Table 5
An example of the original nursery rhyme verse presented in the learning phase (syllables with word stress in bold) and the modified verse presented in the testing phase (acoustic changes in italics, word changes underlined).

Learning phase	Testing phase	Literal translation
Pölhö **Pek**ka **Pöl**lölästä **tor**tun **nä**ki **ker**ran.	***Pöl***(f0 +)hö **Pak**ka **Kok**kolasta ***Tor***(int +)tun **nä**ki **ker**ran.	Simple Pekka from Simpleton saw a pie once.
kysyi **sil**loin **lei**purilta: ”**Sais**kos **pa**lan **ver**ran?”	**kä**syi ***sil***(f0-)loin ***lei***(int-)purilta: ”**Sais**kos **ki**lon **ver**ta?”	He asked from the pie maker: “Could I get a slice?”
”**On**kos **ra**haa, **mil**lä **mak**saa?” **tuu**mi **uk**ko **sii**tä.	”**On**kos **ru**haa, **jot**ta ***mak***(int +)saa?” ***tuu***(f0-)mi **uk**ko **sii**tä.	“Do you have money to pay with”, asked the pie maker.
Pekka **sa**noi:” **Ei**pä **kos**kaan **mul**la **ra**hat **rii**tä.”	**Poi**ka *sa*(f0 +)noi:” **äi**pä **kos**kaan ***mul***(int-)la **ra**hat **rii**ran.”	Pekka said: “My money is never enough.”

f0 + /- = frequency change up/down; int + /- = intensity change up/down.

Infants were asleep in a crib during the measurement, facing randomly either to the left or to the right, thus partly obscuring one ear. The infants can be measured when they are asleep, as the elicitation of prediction error responses does not require focused attention (Cheour et al., 2002). Infants' hearing was normal according to the clinical routine screening done with otoacoustic emission test (EOAE, ILO88, Dpi, Otodynamics Ltd, Hatfield, UK). The measurement was conducted by a registered nurse, who observed the infant and their apparent sleep stages throughout the EEG recording. The stimuli were presented via two loudspeakers located close to the corners at the foot of the crib, the approximate sound level being 74 dB SPL.

2.4. EEG recording and data analysis

EEG was recorded from 12 channels using Ag/Cl-electrodes placed on the infants' head according to the international 10/20 system (NeuroScan, Synamps 2 amplifier). Electrodes were attached to F3, F4, C3, Cz, C4, P3, P4, T7, T8, the two mastoids, and below the right eye. The sampling frequency was 500 Hz and the left mastoid was used as a reference.

The data were re-referenced to the average of the two mastoids and bandpass-filtered (high-pass 0.5 Hz, low-pass 30 Hz, slope 48 dB/octave). Possible eye movements were corrected, and data were divided into epochs of −100-500 ms with respect to the beginnings of the syllables of interest. Epochs with artefacts exceeding 100 μV were rejected and omitted from further analysis. Epochs were then averaged for each stimulus type with a 100 ms pre-stimulus baseline.

As the paradigm included many kinds of deviants, there were on average 60.7 accepted epochs per deviant. Therefore, some of the types had to be grouped to make sure there was a sufficient number of epochs per infant for averaging. The grouping was based on whether the changes were phonological in the Finnish language or not: word or phoneme changes are phonological in Finnish but fundamental frequency and intensity changes are not. The phonological status was expected to be important because the infants born to Finnish-speaking families[1] must have been exposed to the Finnish phonology before birth. Based on previous findings on fetal learning (Partanen et al., 2013b, Partanen, Kujala, Tervaniemi et al., 2013), this might favour the processing of native phonological features over non-phonological ones. Therefore responses to frequency and intensity changes were grouped into the category of acoustic feature changes and word or phoneme changes were grouped into the category of word changes (see also Table 6)[2] .

Difference waveforms were calculated by subtracting the standard ERP (i.e. responses to unaltered syllables in learning phase) from the deviant ERP (i.e., responses to deviations) for each deviant type and each infant in each condition (nursery rhyme, song, prose). Using difference waveforms allows assessing the responses to deviants with minimal contribution of exogenous ERP components related to sound properties. Two time-windows (40–100 ms and 130–190 ms) were selected for statistical analyses based on earlier literature (especially Ylinen et al., 2017) and peaks revealed by visual inspection of grand average difference waves.

2.5. Statistical analysis

The presence of prediction error responses to changes in words and acoustic features in the selected time windows was examined by averaging together the mean amplitudes from F3, F4, C3, C4, Cz, P3 and P4. The averages were submitted to one sample *t*-test in order to assess whether the responses differed significantly from zero. The effect sizes for the t-tests were calculated using Cohen's d values. The normality of the data was inspected visually from histograms depicting standardized residuals.

The differences between different conditions were examined by submitting the mean amplitude to a 3 × 3 x 2 repeated-measures ANOVA with condition (nursery rhyme, song, prose), frontality (electrode lines F, C ja P), and laterality (electrode lines 3 ja 4) as factors. ANOVAs were run separately for each time window. When sphericity could not be assumed, Greenhouse-Geisser correction was used. Statistically significant ($p < .05$) main influences were analyzed in more detail with Bonferroni-corrected pair-wise comparisons. Cohen's d values were calculated for each pair-wise comparison.

[1] At least the mother had to be a native Finnish speaker.
[2] This distinction is not to say that responses to word changes were not acoustically determined. Word changes were phonological but were cued by acoustics. Word meanings, in turn, could not contribute to the processing in newborn infants.

Table 6
Average amplitudes derived from difference waveforms (in μV) to intensity, frequency, word and vowel changes in each condition. The two bars for each electrode illustrate the amplitude in the two time windows (40–100 ms and 130–190 ms). Since the responses for intensity and frequency were similar, they were grouped into the category of acoustic feature changes (not phonological in the Finnish language). Similarly word or phoneme changes were grouped into the category of word changes (phonological in Finnish).

3. Results

Statistically significant prediction error responses to word changes were found in newborns when deviations were presented within the nursery rhyme [40–100 ms: t(20) = −3.32, p = .01; 130–190 ms: t(20) = −2.35, p = .03] (Table 7). Word changes to deviations within song or prose did not elicit responses that differed significantly from zero, and neither did acoustic changes in any

Table 7
t-test for difference waves averaged from electrodes F3, F4, C3, C4, Cz, P3 and P4 (WC = word change, AC = acoustic feature change) in time windows 40–100 ms ja.130–190 ms.

Difference wave	Mean (SD)	t(dF)	*p* value	Cohen's d
	Word change	**40-100 ms**		
Nursery rhyme	−1.09 (1.51)	−3.32 (20)	.01**	0.72
Song	0.07 (1.14)	0.26 (20)	.80	0.06
Speech	−0.23 (0.86)	−1.21 (20)	.24	0.26
	Word change	**130-190 ms**		
Nursery rhyme	−1.23 (2.39)	−2.35 (20)	.03*	0.51
Song	0.18 (1.74)	0.47 (20)	.64	0.10
Speech	−0.28 (0.99)	−1.30 (20)	.21	0.28
	Acoustic change	**40-100 ms**		
Nursery rhyme	0.01 (1.58)	0.021 (20)	.98	0.00
Song	0.19 (1.51)	0.58 (20)	.57	0.13
Speech	−0.28 (1.18)	−1.07 (20)	.30	0.23
	Acoustic change	**130-190 ms**		
Nursery rhyme	0.23 (2.21)	0.47 (20)	.64	0.10
Song	0.48 (1.91)	1.14 (20)	.27	0.25
Speech	−0.06 (1.61)	−0.18 (20)	.86	0.04

** $p < .01$; * $p < .05$.

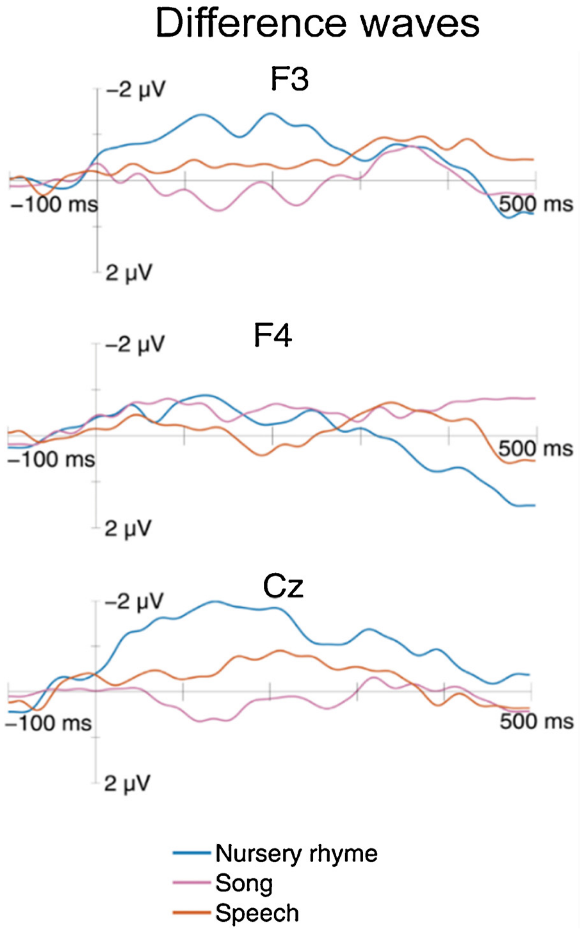

Fig. 4. The brain responses of newborn infants to word changes in stressed syllables from electrodes Cz, F3, and F4.

stimulus type (Figs. 4 and 5).

The ANOVA of the word change responses and their spatial distributions showed a statistically significant main effect of condition [$F(2,40) = 4.99$, $p = .02$, $\eta_p^2 = 0.2$] in the earlier time frame (40–100 ms). In addition, pairwise comparisons showed that the prediction error responses to deviations within the nursery rhyme were more negative than those in the song ($p < .05$, $d = 0.7$). The difference between responses to deviations in the nursery rhyme and in the speech condition was marginally significant ($p = .055$). ANOVA showed no significant differences for acoustic feature changes [$F(2,40) = 0.42$, $p = .66$] in the earlier time window, or for any kind of change in latter time window (word change: [$F(2,40) = 2.98$, $p = .06$], acoustic change: [$F(2,40) = 0.37$, $p = .70$]).

4. Discussion

The current study examined whether the facilitative effect of music and rhythm on learning from auditory input can be observed in newborn infants. Infants were presented with three recorded versions of a nursery rhyme: one spoken metrically as a nursery rhyme, one sung to a melody, and one read as ordinary speech. These naturalistic conditions differed by their duration, intensity, pitch and rhythmic structure. We introduced word deviations and acoustical feature deviations in each version, and expected the changes to elicit prediction error responses, should the infants be able to extract predictions from the recently learned unaltered rendition. Only the word changes in nursery rhyme condition evoked significant prediction error responses. The responses to deviations in the nursery rhyme condition were significantly more negative than those in the song condition and were marginally more negative than those in the speech condition. Acoustic feature changes were not found to elicit prediction error responses in any condition.

These results suggest that the rhythmic structure of nursery rhymes may facilitate learning from auditory input in newborn infants and may thus help future language development. Rhythm likely acts as a framework that allows the brain to form predictions of the future input, and aids temporal sequencing and segmentation (Huss et al., 2011; Kotz, Schwartze, & Schmidt-Kassow, 2009; Schön & Tillmann, 2015). The finding is supported by everyday-life experience: nursery rhymes are recited to and with babies and toddlers in many cultures. Perhaps nursery rhymes have always acted as naturally optimized educational auditory input to infants: their rhythm matches their linguistic prosody, which has been found to enhance phonological processing (Schön & Tillmann, 2015), extraction of

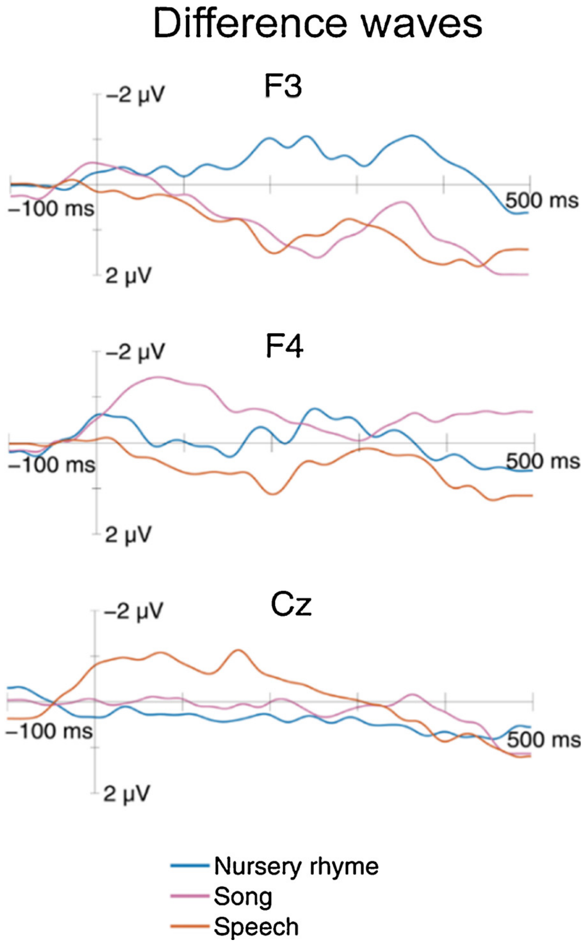

Fig. 5. The brain responses of newborn infants to acoustic changes in stressed syllables from electrodes Cz, F3, and F4.

phonological structure (Leong & Goswami, 2015) and lyric comprehension (Gordon, Magne, & Large, 2011). In addition, the auditory-visual synchrony present when an adult is reciting nursery rhymes promotes heightened attention and memory in infants (Bahrick & Lickliter, 2000; Lewkowicz & Lickliter, 2013; Lewkowicz, 2000).

Previous research with adults suggests that both music and rhythm may help learning by creating templates for future events. Surprisingly, no such effect was found for the song condition in the current study, despite numerous studies that have found music to be beneficial for language learning and word segmentation (e.g. François et al., 2017; Kraus & Chandrasekaran, 2010; Putkinen, Tervaniemi, & Huotilainen, 2013; Putkinen, Saarikivi, & Tervaniemi, 2013; Tallal & Gaab, 2006). There are several possible accounts for this finding. First of all, it has been argued that melody may facilitate learning by attracting, maintaining and enhancing attention. For example, Nakata and Trehub (2004) hypothesized that short fixations associated with maternal speech might be linked to efficient learning and information transfer, while long fixations associated with maternal singing might be optimized to enhance interpersonal ties. Since the participants in the current study were asleep and thus not attending to the stimuli, the facilitating effect of the melody may have been weaker. Secondly, the lack of learning might have been due to a ceiling effect described by Thiessen & Saffran, 2004: single salient prosody cue might have the same effect as multiple cues. This possibility was speculated by Sambeth et al. (2008), who did not find a difference in infants' responses when comparing singing and continuous speech. However, this would not explain why the nursery rhyme and the song conditions did not have similar effect. Thirdly, learning both the segmental (phonetic) and supra-segmental (melodic and/or rhythmic) patterns simultaneously from multi-dimensional signal such as a lyrical song might be more difficult for the newborn infants than learning only the segmental patterns. A previous study by Lebedeva and Kuhl (2010) found that melody facilitated phonetic recognition in 10-month-olds, but preliminary tests in 8-month-olds showed no difference between melodies or spoken strings. Music does contain additional information compared to nursery rhyme, so it is feasible that infants would use some of their cognitive resources to learn the melodic patterns when mapping of linguistic and musical information is not consistent, as in our song condition. Finally, the pattern of findings may suggest that stronger rhythmic structure of the nursery rhyme condition helps infants with probabilistic inference of words, and thus facilitates learning. Younger infants appear to segment speech stimuli based on probabilities and start to prefer prosodic cues only around eight months of age; possibly because they start to lose sensitivity to non-native contrasts (Johnson & Jusczyk, 2001; Thiessen & Saffran, 2007; Yeung & Werker, 2009).

Our results support the previous findings from infants suggesting that exaggerated or variable pitch information within signal can facilitate phonetic recognition (Bergeson & Trehub, 2007; Lebedeva & Kuhl, 2010; Thiessen, Hill, & Saffran, 2005). As Tables 3 and 4 and Figs. 2 and 3 show, the current nursery rhyme stimuli, which led to the successful detection of word changes, had more pitch variation than the song or speech stimuli. It seems that pitch modulations might facilitate word segmentation as François et al. (2017) suggested, but here they were more pronounced in the nursery rhyme than the song. The difference in our and François' and colleagues' results might be due to us using a simple melody, that did not match well with word-stress pattern, whereas they had associated each syllable to a unique pitch. In line with François et al., we suggest that an exaggerated prosodic cue, such as pitch change, increases the saliency of to-be-learned item. What the present results add to those by François and colleagues is that the regular rhythmic patterns of the input, as introduced by the current nursery rhyme, seem to synchronize the auditory processing and to facilitate the prediction of the following items. This would also be in line with Goswami (2018) oscillatory temporal sampling framework. Perhaps the stronger frequency modulation and periodic power in the stressed syllable modulation tier in nursery rhyme stimuli lead to more efficient phase entrainment of neuronal oscillations in the auditory cortex, and thus more accurate temporal sampling of stressed syllables and more efficient learning. The song condition had higher periodic power in syllable modulation tier, which also suggests that while there was rhythmic regularity in the song stimulus, the stressed syllables were not as salient in the song condition as in the nursery rhyme condition.

The present study is one of a few in which infants were presented with continuous speech. Thus, the natural stimuli are highly complex and acoustically not very well controlled. As follows, the current findings should be considered with some caution. When interpreting the results, it is good to keep in mind that infants' brain responses are characterized by remarkable variability, the reasons of which are not yet fully understood. All conditions elicited positive responses in some infants, and negative responses in some other infants. The same phenomenon has been previously observed in traditional oddball paradigms (reviewed in Näätänen, Sussman, Salisbury, & Shafer, 2014) and also when infants have been exposed to the statistical auditory streams (Bosseler, Teinonen, Tervaniemi, & Huotilainen, 2016; Teinonen et al., 2009). In some conditions, these different response patterns may have cancelled each other out in averaging. Such variability might account for not observing significant responses to acoustic changes that typically elicit quite robust responses. In addition, the lack of significant responses may be caused by the use of continuous stimuli, the learning of which is expected to be more challenging than that of traditional oddball streams with simple acoustic deviations. Given the promising results, it would be valuable to replicate this study with a larger participant sample and continuous but more controlled stimuli.

In the future, it would also be interesting to complement this study by exploring which factors in music and rhythm promote learning across different developmental stages. It might be that the effects of music and rhythms interact with effective attention allocation, and prediction errors would be associated with redirecting attention (Ylinen et al., 2017). Alternatively, language experience and developmental changes might be required before an infant can differentially attend to phonetic and melody information.

5. Conclusion

Our findings suggest that rhythmic structure of nursery rhymes may facilitate newborn infants' learning from auditory input and may thus be beneficial to language development. Surprisingly, no such effect was found for a song in the present stimuli. This might be due to melody not including enough pitch variation, or simultaneous learning of the melody and phonetic content being more challenging to newborn infants than that of the phonetic content alone. In line with previous behavioral IDS studies (Curtin, Campbell, & Hufnagle, 2012; Spinelli, Fasolo, & Mesman, 2017), it seems that simultaneous occurrence of exaggerated prosodic cues and to-be-learned items helps infants to process language input. This could be taken into consideration for example when planning activities and interactions at home, music playschools, or when making educational music: linguistic and musical rhythm should maintain synchrony.

Author contributions statement

S.Y. and M.H. designed research; S.Y. and E.S. performed research, E.S. analyzed data; and E.S., M.H., and S.Y. wrote the paper.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

The authors would like to thank all the families for participating, Tarja Ilkka for conducting the EEG measurements, and Tanja Linnavalli and Valtteri Wikström for composing the song and recording the stimuli. We would also like to thank Tommi Makkonen and Miika Leminen for their technical assistance. This project is conducted by University of Helsinki and Hospital District of Helsinki and Uusimaa and supported by Research Funds of the University of Helsinki and the Academy of Finland [Grants 276414 and 131963].

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.infbeh.2019.101346.

References

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*(2), 111–130. https://doi.org/10.1016/S0022-0965(02)00124-8.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351–373. https://doi.org/10.1016/j.wocn.2012.02.003.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*(2), 190. https://doi.org/10.1037//0012-1649.36.2.190.

Bergeson, T. R., & Trehub, S. E. (2007). Signature tunes in mothers' speech to infants. *Infant Behavior & Development, 30*(4), 648–654. https://doi.org/10.1016/j.infbeh.2007.03.003.

Bergman Nutley, S., Darki, F., & Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Frontiers in Human Neuroscience, 7*, 926. https://doi.org/10.3389/fnhum.2013.00926.

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience, 25*(3-4), 399–410.

Bhatara, A., Yeung, H. H., & Nazzi, T. (2015). Foreign language learning in French speakers is associated with rhythm perception, but not with melody perception. *Journal of Experimental Psychology Human Perception and Performance, 41*(2), 277. https://doi.org/10.1037/a0038736.

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer (version 6.0.40) [computer program].* Retrieved June 4, 2018 fromhttp://www.praat.org/.

Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PloS One, 11*(9), e0162177. https://doi.org/10.1371/journal.pone.0162177.

Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., et al. (2002). Psychobiology: Speech sounds learned by sleeping newborns. *Nature, 415*(6872), 599. https://doi.org/10.1038/415599b.

Chobert, J., François, C., Velay, J. L., & Besson, M. (2012). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex (New York, NY: 1991), 24*(4), 956–967. https://doi.org/10.1093/cercor/bhs377.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science, 17*(6), 1003–1011. https://doi.org/10.1111/desc.12193.

Corrigall, K. A., & Trainor, L. J. (2011). Associations between length of music training and reading skills in children. *Music Perception, 29*(2), 147–155. https://doi.org/10.1525/mp.2011.29.2.147.

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition, 96*(3), 233–262. https://doi.org/10.1016/j.cognition.2004.08.005.

Curtin, S., Campbell, J., & Hufnagle, D. (2012). Mapping novel labels to actions: How the rhythm of words guides infants' learning. *Journal of Experimental Child Psychology, 112*(2), 127–140. https://doi.org/10.1016/j.jecp.2012.02.007.

Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. *Proceedings of the international congress of phonetics science,* 471–474 ISBN 1-876346-48-5.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117. https://doi.org/10.1038/nn1504.

Emberson, L. L., Richards, J. E., & Aslin, R. N. (2015). Top-down modulation in the infant brain: Learning-induced expectations rapidly affect the sensory cortex at 6 months. In: *Proceedings of the National Academy of Sciences of the United States of America, 112*(31), 9585–9590. https://doi.org/10.1073/pnas.1510343112.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infantpreference for motherese speech. *Infant Behavior & Development, 10*, 279–293. https://doi.org/10.1016/0163-6383(87)90017-8.

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., et al. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience, 8*, 392. https://doi.org/10.3389/fnhum.2014.00392.

François, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex (New York, NY: 1991), 21*(10), 2357–2365. https://doi.org/10.1093/cercor/bhr022.

François, C., Chobert, J., Besson, M., & Schön, D. (2012). Music training for the development of speech segmentation. *Cerebral Cortex (New York, NY: 1991), 23*(9), 2038–2043. https://doi.org/10.1093/cercor/bhs180.

François, C., Teixidó, M., Takerkart, S., Agut, T., Bosch, L., & Rodriguez-Fornells, A. (2017). Enhanced neonatal brain responses to sung streams predict vocabulary outcomes by age 18 months. *Scientific Reports, 7*(1), 12451. https://doi.org/10.1038/s41598-017-12798-2.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris, 100*(1-3), 70–87. https://doi.org/10.1016/j.jphysparis.2006.10.001.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138. https://doi.org/10.1038/nrn2787.

Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PloS One, 3*(10), e3566. https://doi.org/10.1371/journal.pone.0003566.

Gagnepain, P., Henson, R. N., & Davis, M. H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Current Biology: CB, 22*(7), 615–621. https://doi.org/10.1016/j.cub.2012.02.015.

Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Developmental Science, 15*(3), 398–407.

Gordon, R. L., Magne, C. L., & Large, E. W. (2011). EEG correlates of song prosody: A new look at the relationship between linguistic and musical rhythm. *Frontiers in Psychology, 2*, 352. https://doi.org/10.3389/fpsyg.2011.00352.

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & McAuley, J. D. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science, 18*(4), 635–644. https://doi.org/10.1111/desc.12230.

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. In: *Proceedings of the National Academy of Sciences of the United States of America, 99*(16), 10911–10916. https://doi.org/10.1073/pnas.122368599.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, 15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001.

Goswami, U., Wang, H. L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience, 23*(2), 325–337. https://doi.org/10.1162/jocn.2010.21453.

Goswami, U. (2018). A neural basis for phonological awareness? An oscillatory temporal-sampling perspective. *Current Directions in Psychological Science, 27*(1), 56–63. https://doi.org/10.1177/0963721417727520.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In N. Warner, & C. Gussenhoven (Vol. Eds.), *Papers in laboratory phonology: vol. 7*, (pp. 515–546). Berlin, Germany: Mouton de Gruyter.

Graniere-Deferre, C., Ribeiro, A., Jacquet, A. Y., & Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Developmental Science, 14*(2), 336–352. https://doi.org/10.1111/j.1467-7687.2010.00978.x.
Háden, G. P., Németh, R., Török, M., & Winkler, I. (2015). Predictive processing of pitch trends in newborn infants. *Brain Research, 1626*, 14–20. https://doi.org/10.1016/j.brainres.2015.02.048.
Ho, Y. C., Cheung, M. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology, 17*(3), 439. https://doi.org/10.1037/0894-4105.17.3.439.
Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex, 47*(6), 674–689. https://doi.org/10.1016/j.cortex.2010.07.010.
Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*(4), 548–567. https://doi.org/10.1006/jmla.2000.2755.
Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323. https://doi.org/10.1037/0033-295X.83.5.323.
Kayhan, E., Meyer, M., O'Reilly, J. X., Hunnius, S., & Bekkering, H. (2019). Nine-month-old infants update their predictive models of a changing environment. *Developmental Cognitive Neuroscience*100680. https://doi.org/10.1016/j.dcn.2019.100680.
Kivijärvi, M., Voeten, M. J., Niemelä, P., Räihä, H., Lertola, K., & Piha, J. (2001). Maternal sensitivity behavior and infant behavior in early interaction. *Infant Mental Health Journal, 22*(6), 627–640. https://doi.org/10.1002/imhj.1023.
Kouider, S., Long, B., Le Stanc, L., Charron, S., Fievet, A. C., Barbosa, L. S., et al. (2015). Neural dynamics of prediction and surprise in infants. *Nature Communications, 6*, 8537. https://doi.org/10.1038/ncomms9537.
Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599. https://doi.org/10.1038/nrn2882.
Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., et al. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience, 34*(36), 11913–11918. https://doi.org/10.1523/JNEUROSCI.1881-14.2014.
Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex, 45*(8), 982–990. https://doi.org/10.1016/j.cortex.2009.02.010.
Kunnas, K. (1997). *Hanhiemon iloinen lipas.* Helsinki: WSOY.
Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: Infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant Behavior & Development, 33*(4), 419–430. https://doi.org/10.1016/j.infbeh.2010.04.006.
Lecanuet, J. P., Graniere-Deferre, C., Jacquet, A. Y., & DeCasper, A. J. (2000). Fetal discrimination of low-pitched musical notes. *Developmental Psychobiology, 36*(1), 29–39. https://doi.org/10.1002/(SICI)1098-2302(200001)36:1%3C29::AID-DEV4%3E3.0.CO;2-J.
Lee, C. S., Kitamura, C., Burnham, D., & Todd, N. P. (2014). On the rhythm of infant-versus adult-directed speech in Australian English. *The Journal of the Acoustical Society of America, 136*(1), 357–365. https://doi.org/10.1121/1.4883479.
Leong, V. (2012). *Prosodic rhythm in the speech amplitude envelope: Amplitude modulation phase hierarchies (AMPHs) and AMPH models*Cambridge, UK: PhD Thesis.
Leong, V., Kalashnikova, M., Burnham, D., & Goswami, U. (2014). Infant-directed speech enhances temporal rhythmic structure in the envelope. *InterSpeech 2014, proceedings of the 15th Annual Conference of the International Speech Communication Association (ISCA)*.
Leong, V., & Goswami, U. (2015). Acoustic-emergent phonology in the amplitude envelope of child-directed speech. *PloS One, 10*(12), e0144411. https://doi.org/10.1371/journal.pone.0144411.
Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin, 126*(2), 281.
Lewkowicz, D. J., & Lickliter, R. (2013). *The development of intersensory perception: Comparative perspectives.* Hillsdale, NJ: Psychology Press.
Linnavalli, T., Putkinen, V., Lipsanen, J., Huotilainen, M., & Tervaniemi, M. (2018). Music playschool enhances children's linguistic skills. *Scientific Reports, 8*(1), 8767. https://doi.org/10.1038/s41598-018-27126-5.
Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*(2), 199–211. https://doi.org/10.1162/jocn.2006.18.2.199.
Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing, 26*(5), 739–769. https://doi.org/10.1007/s11145-012-9389-0.
Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior & Development, 27*(4), 455–464. https://doi.org/10.1016/j.infbeh.2004.03.002.
Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology Human Perception and Performance, 24*(3), 756. https://doi.org/10.1037/0096-1523.24.3.756.
Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 115*(1), 140–144. https://doi.org/10.1016/j.clinph.2003.04.001.
Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026.
Näätänen, R., Sussman, E. S., Salisbury, D., & Shafer, V. L. (2014). Mismatch negativity (MMN) as an index of cognitive dysfunction. *Brain Topography, 27*(4), 451–466. https://doi.org/10.1007/s10548-014-0374-6.
Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., et al. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology, 4*, 10. https://doi.org/10.3389/fpsyg.2013.00010.
Overy, K. (2003). Dyslexia and music. *Ann. NY Acad. Sci. 999*(1), 497–505. https://doi.org/10.1196/annals.1284.060.
Papousek, H. (1996). *Musicality in infancy research: Biological and cultural origins of early musicality. Musical beginnings: Origins and development of musical competence.* Oxford: Oxford University Press37–55.
Parkkonen, L., Andersson, J., Hämäläinen, M., & Hari, R. (2008). Early visual brain areas reflect the percept of an ambiguous scene. In: *Proceedings of the National Academy of Sciences of the United States of America, 105*(51), 20500–20504. https://doi.org/10.1073/pnas.0810966105.
Partanen, E., Kujala, T., Tervaniemi, M., & Huotilainen, M. (2013). Prenatal music exposure induces long-term neural effects. *PloS One, 8*(10), e78946. https://doi.org/10.1371/journal.pone.0078946.
Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. In: *Proceedings of the National Academy of Sciences of the United States of America, 110*(37), 15145–15150. https://doi.org/10.1073/pnas.1302159110.
Peterson, D. A., & Thaut, M. (2007). Music increases frontal EEG coherence during verbal learning. *Neuroscience Letters, 412*(3), 217–221. https://doi.org/10.1016/j.neulet.2006.10.057.
Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *The Behavioral and Brain Sciences, 36*(4), 329–347. https://doi.org/10.1017/S0140525X12001495.
Purnell-Webb, P., & Speelman, C. P. (2008). Effects of music on memory for text. *Perceptual and Motor Skills, 106*(3), 927–957. https://doi.org/10.2466/pms.106.3.927-957.
Putkinen, V., Tervaniemi, M., & Huotilainen, M. (2013). Informal musical activities are linked to auditory discrimination and attention in 2–3-year-old children: An event-related potential study. *The European Journal of Neuroscience, 37*(4), 654–661. https://doi.org/10.1111/ejn.12049.
Putkinen, V. J., Saarikivi, K. A., & Tervaniemi, M. (2013). Do informal musical activities shape auditory skill development in preschool-age children? *Frontiers in Psychology, 4*, 572. https://doi.org/10.3389/fpsyg.2013.00572.
Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America, 105*(1), 512–521. https://doi.org/10.1121/1.424522.
Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265–292. https://doi.org/10.1016/S0010-0277(99)

00058-X.
Rao, R., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87. https://doi.org/10.1038/4580.
Reynolds, G. D., Bahrick, L. E., Lickliter, R., & Guy, M. W. (2014). Neural correlates of intersensory processing in 5-month-old infants. *Developmental Psychobiology, 56*(3), 355–372. https://doi.org/10.1002/dev.21104.
Rothermich, K., Schmidt-Kassow, M., Schwartze, M., & Kotz, S. A. (2010). Event-related potential responses to metric violations: Rules versus meaning. *Neuroreport, 21*(8), 580–584. https://doi.org/10.1097/WNR.0b013e32833a7da7.
Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia, 50*(2), 232–244. https://doi.org/10.1016/j.neuropsychologia.2011.10.025.
Ruusuvirta, T., Huotilainen, M., Fellman, V., & Näätänen, R. (2009). Numerical discrimination in newborn infants as revealed by event-related potentials to tone sequences. *The European Journal of Neuroscience, 30*(8), 1620–1624. https://doi.org/10.1111/j.1460-9568.2009.06938.x.
Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 119*(2), 332–341. https://doi.org/10.1016/j.clinph.2007.09.144.
Schmidt-Kassow, M., & Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience, 21*(9), 1693–1708. https://doi.org/10.1162/jocn.2008.21153.
Schmuckler, M. A., & Boltz, M. G. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception & Psychophysics, 56*(3), 313–325. https://doi.org/10.3758/BF03209765.
Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition, 106*(2), 975–983. https://doi.org/10.1016/j.cognition.2007.03.005.
Schön, D., & Tillmann, B. (2015). Short-and long-term rhythmic interventions: Perspectives for language rehabilitation. *Annals of the New York Academy of Sciences, 1337*(1), 32–39. https://doi.org/10.1111/nyas.12635.
Shannon, K. (2006). Infant behavioral responses to infant-directed singing and other maternal interactions. *Infant Behavior & Development, 29*(3), 456–470. https://doi.org/10.1016/j.infbeh.2006.03.002.
Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review: DR, 44*, 1–18. https://doi.org/10.1016/j.dr.2016.12.001.
Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences, 29*(7), 382–390. https://doi.org/10.1016/j.tins.2006.06.003.
Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience, 10*(1), 21. https://doi.org/10.1186/1471-2202-10-21.
Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., & Wartenburger, I. (2011). Acoustic processing of temporally modulated sounds in infants: Evidence from a combined near-infrared spectroscopy and EEG study. *Frontiers in Psychology, 2*, 62. https://doi.org/10.3389/fpsyg.2011.00062.
Thaut, M. (2005). *Rhythm, music, and the brain: Scientific foundations and clinical applications.* New York: Routledge.
Thiessen, E. D., & Saffran, J. R. (2004). Spectral tilt as a cue to word segmentation in infancy and adulthood. *Perception & Psychophysics, 66*(5), 779–791. https://doi.org/10.3758/BF03194972.
Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5.
Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development : the Official Journal of the Society for Language Development, 3*(1), 73–100. https://doi.org/10.1080/15475440709337001.
Torppa, R., Faulkner, A., Huotilainen, M., Järvikivi, J., Lipsanen, J., Laasonen, M., et al. (2014). The perception of prosody and associated auditory cues in early-implanted children: The role of auditory working memory and musical activities. *International Journal of Audiology, 53*(3), 182–191. https://doi.org/10.3109/14992027.2013.872302.
Trainor, L. J., Lee, K., & Bosnyak, D. J. (2011). Cortical plasticity in 4-month-old infants: Specific effects of experience with musical timbres. *Brain Topography, 24*(3-4), 192. https://doi.org/10.1007/s10548-011-0177-y.
Trainor, L. J. (2012). Predictive information processing is a fundamental learning mechanism present in early development: Evidence from infants. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology, 83*(2), 256–258. https://doi.org/10.1016/j.ijpsycho.2011.12.008.
Trehub, S. E., & Schellenberg, E. G. (1995). Music: Its relevance to infants. *Ann. Child Dev. 11*, 1–24.
Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*(1), 73–99. https://doi.org/10.1016/j.cognition.2005.11.006.
Virtala, P., & Partanen, E. (2018). Can very early music interventions promote at-risk infants' development? *Annals of the New York Academy of Sciences, 1423*(1), 92–101. https://doi.org/10.1111/nyas.13646.
Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2001). *The origins of music*. MIT press.
White, L., & Mattys, S. L. (2007a). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.). *Segmental and prosodic issues in romance phonology* (pp. 237–257). Amsterdam: John Benjamins.
White, L., & Mattys, S. L. (2007b). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35*(4), 501–522. https://doi.org/10.1016/j.wocn.2007.02.003.
Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. (2010). How stable are acoustic metrics of contrastive speech rhythm? *The Journal of the Acoustical Society of America, 127*, 1559–1569. https://doi.org/10.1121/1.3293004.
Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology, 21*(3-4), 147–163. https://doi.org/10.1027/0269-8803.21.34.147.
Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. In: *Proceedings of the National Academy of Sciences of the United States of America, 106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106.
Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition, 113*(2), 234–243. https://doi.org/10.1016/j.cognition.2009.08.010.
Ylinen, S., Huuskonen, M., Mikkola, K., Saure, E., Sinkkonen, T., & Paavilainen, P. (2016). Predictive coding of phonological rules in auditory cortex: A mismatch negativity study. *Brain and Language, 162*, 72–80. https://doi.org/10.1016/j.bandl.2016.08.007.
Ylinen, S., Bosseler, A., Junttila, K., & Huotilainen, M. (2017). Predictive coding accelerates word recognition and learning in the early stages of language development. *Developmental Science, 20*(6), https://doi.org/10.1111/desc.12472.
Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. In: *Proceedings of the National Academy of Sciences of the United States of America, 113*(19), 5212–5217. https://doi.org/10.1073/pnas.1603984113.